

AMPLITUDE-SCALING RESILIENT AUDIO WATERMARKING METHOD AND APPARATUS BASED ON QUANTIZATION

BACKGROUND OF THE INVENTION

5

Field of the Invention

[0001] The present invention relates to an audio watermarking apparatus and method, and more particularly, to an amplitude-scaling resilient audio watermarking method
10 based on a quantization.

Discussion of the Related Art

[0002] Recently, illegal distribution of the digital audio contents over the Internet occurs frequently. Therefore,
15 apparatuses for the copyright protection of digital audio contents are required. Audio watermarking is a method for copyright protection through embedding copyright information in digital audio contents. Embedded watermark should be imperceptible and robust against signal processing procedures
20 and malicious attacks.

[0003] LSB modulation, phase shift keying, echo hiding, spread spectrum watermarking, and quantization watermarking have been proposed as audio watermarking methods.

[0004] Watermarking method can be categorized as the blind
25 watermarking and the non-blind watermarking with respect to its decoding scheme. The blind watermarking method decodes

the embedded watermark without access to the host signal, in which a watermark is not embedded. Early blind watermarking methods are based on the spread spectrum technique, which reduces the host-signal interference by employing a modulation scheme with a long pseudorandom sequence. Also, An advanced quantization watermarking method, which employs the side information at the encoder, has been proposed. In comparison with the conventional spread spectrum watermarking, the advanced quantization watermarking provides better performance by reducing the host-signal interference in the detection process.

[0005] However, the quantization watermarking is vulnerable to the amplitude scaling. In other words, if the amplitude of the watermarked signal is scaled by a constant ratio, the decoding performance may be degraded greatly by the mismatch between the amplitude of the decoder's input signal and the quantizer step size of the decoder.

[0006] U.S. Pat. No. 6,483,927 discloses a watermarking method based on a quantization, which compensates the attack distortion by estimating the applied attack. In the patent, the embedding region may be determined as the amplitude of the signal, or the transformation coefficients such as the coefficients of DCT, DWT, DFT and the like.

[0007] J. J. Eggers, R. Bäuml, R. Tzschoppe and B. Girod, "Scalar Costa Scheme for Information Embedding," IEEE Transactions on Signal Processing, vol. 51, No. 4, April 2003,

pp. 1003-1019, discloses a Scalar Costa Scheme (SCS) for embedding and decoding a watermark using a codebook, which is constructed using uniform scalar quantizers.

[0008] The Scalar Costa Scheme (SCS) is a blind
5 watermarking method, which reduces the host-signal interference, and it employs the uniform scalar quantizer for practical implementation. Although watermarking method, which employs the uniform scalar quantizer, is practical with simple implementation, it is very vulnerable to the amplitude
10 scaling, which modifies the amplitude of the watermarked signal.

[0009] Accordingly, for the purpose of reliable detection, the quantizer step size of the decoder should be adjusted according to the applied amplitude scaling. The conventional
15 decoder performs the decoding process without adjusting the quantizer step size, thus causing a serious degradation of decoding performance. Additionally, since the amplitude scaling of the audio signal occurs frequently, the decoding of the watermark from the amplitude-scaled signal should be
20 considered importantly. The normalization of audio signals with respect to the root mean square (RMS) value of the amplitude is an example of the amplitude scaling.

[0010] Additionally, in order to reliably decode the watermark from the amplitude-scaled signal, Eggers, et. al.
25 proposed an algorithm for estimating the scale factor by using the SCS pilot signal. In the proposed algorithm, a

pilot signal is embedded in a manner of the Scalar Costa Scheme (SCS) and the scale factor is estimated through a Fourier analysis of histograms of the pilot.

[0011] However, in the conventional method, the pilot signal should be long enough to accurately estimate the scale factor. Since the total length of the host signal is finite, the space for embedding the payload decreases as the length of the pilot signal increases.

SUMMARY OF THE INVENTION

[0012] Accordingly, the present invention is directed to an audio watermarking method and apparatus based on a quantization that substantially obviates one or more problems due to limitations and disadvantages of the related art.

[0013] An object of the present invention is to provide an audio watermarking apparatus and method based on a quantization, in which the scale factor of the watermarked signal is estimated just before the actual decoding process by using the expectation maximization (EM) algorithm, and the quantizer step size is adjusted, thereby providing an amplitude-scaling resilient decoding result.

[0014] Additional advantages, objects, and features of the invention will be set forth in part in the description which follows and in part will become apparent to those having ordinary skill in the art upon examination of the following or may be learned from practice of the invention. The

objectives and other advantages of the invention may be realized and attained by the structure particularly pointed out in the written description and claims hereof as well as the appended drawings.

5 **[0015]** To achieve these objects and other advantages and in accordance with the purpose of the invention, as embodied and broadly described herein, an amplitude-scaling resilient audio watermarking encoding apparatus based on a quantization includes: a polyphase filterbank for dividing an inputted
10 audio signal into a plurality of subbands; a psychoacoustic module for applying a psychoacoustic model to the inputted audio signal to provide a signal-to-mask ratio (SMR); a watermark encoder for evaluating an encoding parameter from the plurality of subbands according to the signal-to-mask
15 ratio (SMR) provided from the psychoacoustic module and embedding the encoding parameter and a watermark into subbands corresponding to the middle frequency among the plurality of subbands; and a synthesis filterbank for synthesizing the divided and watermarked subband signals to
20 output a watermarked audio signal.

[0016] An amplitude-scaling resilient audio watermarking decoding apparatus based on a quantization includes: a polyphase filterbank for dividing a received audio signal into the predetermined subbands; an expectation maximization
25 (EM) estimator for estimating the scale factor from an encoding parameter contained in the received audio signal and

a watermarked subband according to the EM algorithm, and generating the quantizer step size Δ_d of a decoder according to the scale factor; a watermark decoder for extracting a watermark from the selected subband using the estimated
5 quantizer step size; and an integrated determiner for integrating outputs of the watermark decoder to determine a watermark.

[0017] A method for encoding an audio signal includes the steps of: dividing an inputted audio signal into subbands;
10 applying a psychoacoustic model to the audio signal to evaluate a signal-to-mask ratio (SMR); evaluating an encoding parameter from the signal-to-mask ratio (SMR); encoding a watermark in each subband according to the evaluated encoding parameter; synthesizing the watermarked subbands; and
15 transmitting watermarked audio signal and the encoding parameter.

[0018] A method for decoding an audio signal includes the steps of: receiving the audio signal and a side information; dividing the audio signal into subbands; estimating a scale
20 factor from the side information and the received audio signal by using an expectation maximization (EM) algorithm, and evaluating the quantizer step size of a decoder from the estimated scale factor; decoding a watermark from the subbands using the evaluated quantizer step size; and summing
25 up the decoded values to calculate an average, and

calculating a correlation between the average and each codeword of the codebook to determine the embedded watermark.

[0019] It is to be understood that both the foregoing general description and the following detailed description of the present invention are exemplary and explanatory and are intended to provide further explanation of the invention as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] The accompanying drawings, which are included to provide a further understanding of the invention and are incorporated in and constitute a part of this application, illustrate embodiment(s) of the invention and together with the description serve to explain the principle of the invention. In the drawings:

[0021] FIG. 1 illustrates a concept of the quantization watermarking, which is applied to the present invention;

[0022] FIG. 2 is a block diagram of a watermark encoding apparatus according to the present invention;

[0023] FIG. 3 is a block diagram of the watermark encoder shown in FIG. 2;

[0024] FIG. 4 is a flowchart showing a watermarking encoding method according to the present invention;

[0025] FIG. 5 is a block diagram of a watermarking decoding apparatus according to the present invention;

[0026] FIG. 6 is a flowchart showing a watermarking decoding method according to the present invention; and

[0027] FIG. 7 illustrates simulation results in case that both MP3 lossy compression and amplitude-scaling are applied.

5

DETAILED DESCRIPTION OF THE INVENTION

[0028] Reference will now be made in detail to the preferred embodiments of the present invention, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers will be used throughout the drawings to refer to the same or like parts.

[0029] FIG. 1 illustrates a concept of the quantization watermarking, which will be applied to the present invention.

[0030] Referring to FIG. 1, the quantization watermarking is a method of embedding the watermark by quantizing an audio signal with the quantizer, which is selected according to the corresponding watermark sequence. In other words, the quantization is performed using a quantizer 1 and a quantizer 0, whose quantization reference level is shifted by $\Delta/2$. If a value of a watermark sequence d_n is "1", the quantization is performed by the quantizer 1, and if the value is "0", the quantization is performed by the quantizer 0.

[0031] Meanwhile, the quantization watermarking is vulnerable to the amplitude scaling. When the amplitude-scaling is applied to the watermarked signal, the mismatch between the watermarked signal and the quantizer step size of

the decoder can degrade the decoding performance. According to the present invention, the quantizer step size is adjusted through an estimation of the applied scale factor. The scale factor is estimated from the input signal of the decoder by
5 the expectation maximization (EM) algorithm.

[0032] The present invention employs a blind type detection method and the host signal information at the encoder is exploited in the process of the watermark encoding in order to reduce the host-signal interference. In order for
10 robustness against the lossy compression and the general signal processing, the watermark is repeatedly embedded into the subbands corresponding to the middle frequency. A final result is obtained by integrating each result of the subbands. Since the robustness against attacks varies with respect to
15 the subband, integrating can provide more robustness.

[0033] An audio watermarking system of the present invention is generally divided into an encoding apparatus and a decoding apparatus.

[0034] 1. ENCODING APPARATUS

20 [0035] FIG. 2 is a block diagram of a encoding apparatus according to the present invention, FIG. 3 is a embedding algorithm of the watermark encoder of FIG. 2, and FIG. 4 is a flowchart showing a watermarking encoding method according to the present invention.

25 [0036] Referring to FIG. 2, the encoding apparatus 200 of the present invention includes a polyphase filterbank 210 for

dividing an inputted audio signal x_n into 32 subbands according to frequencies, a psychoacoustic module 220 for applying a psychoacoustic model to the inputted audio signal to provide a signal-to-mask ratio (SMR), a watermark encoder 5 230 for embedding a watermark into middle frequency subbands among the divided subbands according to the signal-to-mask ratio (SMR) of the psychoacoustic module 220 and providing side information, and a synthesis filterbank 240 for synthesizing subband signals to output a watermarked audio 10 signal.

[0037] In the encoding apparatus 200, the inputted audio signal x_n is divided into 32 subbands by the polyphase filterbank 210. In an embodiment of the present invention, considering robustness against the lossy compression, 15 inaudibility and the like, the watermarks are embedded into fourth to nineteenth subbands corresponding to the middle frequency. Since robustness to compression and amplitude scaling is different in each subband according to the corresponding frequencies, the same watermark signal d_n is 20 repeatedly embedded into the 16 subbands.

[0038] For the inaudibility, an intensity of each watermark is determined using the psychoacoustic model. In the 16 watermarked subbands, corresponding encoding parameters Δ_e and α are transmitted to each subband as the 25 side information together with the watermarked audio signal.

Here, Δ_e represents the quantizer step size of an encoder and α represents a scale.

[0039] Referring to FIG. 3, the watermark encoder 230 for embedding the watermark into the host signal x_n with respect to each subband includes: a parameter evaluator 231 for evaluating the encoding parameters Δ_e and α from the signal-to-mask ratio (SMR) provided from the psychoacoustic model and an estimation value (WNR) of a noise intensity determined by a specification of a lossy compression; a quantizer 232 for performing a uniform scalar quantization with respect to the audio signal x_n according to the quantizer step size Δ_e by using a quantizer selected by the watermark d_n ; an adder 233 for subtracting the host signal x_n from an output of the quantizer 232; a multiplier 234 for multiplying an output of the adder 233 by the scale α ; and an adder 235 for adding an output of the multiplier 234 to the host signal x_n to output a watermarked subband signal s_n . The watermark embedding algorithm of the present invention is similar to a watermarking method of Scalar Costa Scheme (SCS) proposed by Eggers et. al.

[0040] The process of embedding the watermark in the encoding apparatus is implemented through a dithered scalar quantizer. For an input x that is a constant, $Q_{\Delta,d}(x)$ is defined by an equation 1.

$$[0041] \quad Q_{\Delta,d}(x) \equiv \Delta \left(\left\lfloor \frac{x}{\Delta} - \frac{d}{2} + \frac{1}{2} \right\rfloor + \frac{d}{2} \right) \quad (\text{Eq. 1})$$

where, $\lfloor c \rfloor$ means a maximum integer that is less than or equal to a real number c , a positive constant Δ represents the quantizer step size, and d represents a dither signal having a binary value.

5 **[0042]** A sequence x_n of real number represents an host signal (an audio signal). A watermark message is expressed with a binary sequence d_n through a pseudorandom sequence. When the sequence s_n of real number represents the watermarked signal, the watermark embedding process is given
10 by an equation 2.

$$\text{[0043]} \quad s_n = (1 - \alpha)x_n + \alpha Q_{\Delta_e, d_n}(x_n) \quad (\text{Eq. 2})$$

[0044] Here, α ($0 < \alpha < 1$) and Δ_e are the encoding parameters used in the embedding process and determined differently according to each subband. The values of the encoding
15 parameters Δ_e and α are determined from the signal-to-mask ratio (SMR) provided from the psychoacoustic model and the estimation value (WNR) of the noise intensity determined by the specification of the lossy compression. These values are transmitted to the decoding apparatus together with the
20 watermarked signal.

[0045] A method for encoding the audio signal in the encoding apparatus is shown in FIG. 4. The encoding method includes the steps of: inputting the audio signal (401); dividing the inputted audio signal into subbands (402);
25 applying a psychoacoustic model to the audio signal to

evaluate a signal-to-mask ratio (SMR) (403); evaluating an encoding parameter from the signal-to-mask ratio (SMR) (404); encoding a watermark in each subband according to the evaluated encoding parameter (405); synthesizing the watermark encoded subbands (406); and transmitting watermarked audio signal and the encoding parameter.

[0046] 2. DECODING APPARATUS

[0047] FIG. 5 is a block diagram of a watermark decoding apparatus according to the present invention, and FIG. 6 is a flowchart showing a watermarking decoding method according to the present invention.

[0048] Referring to FIG. 5, the decoding apparatus 500 of the present invention includes: a polyphase filterbank 510 for dividing a received audio signal into 32 subbands; an expectation maximization (EM) estimator 520 for estimating an scale factor from a received encoding parameter and a watermarked subband according to the EM algorithm, and generating the quantizer step size Δ_d of a decoder according to the amplitude scaling; a watermark decoder 530 for extracting a watermark from the subband corresponding to the middle frequency considering the quantizer step size of the decoder; and an integration determiner 540 for integrating outputs of the watermark decoder 530 to determine the watermark.

[0049] A watermark detection in the decoding apparatus 500 is generally carried out through two processes, i.e., a

process of estimating the amplitude-scaling and a process of integrating the decoded signals. In the same manner described in the encoding apparatus, a rate g' is estimated according to the 32 divided subbands and the estimated rate is used to
 5 adjust the quantizer step size Δ_d to $g'\Delta_e$. The watermark extracted according to the subbands is obtained and a final result is calculated by comparing the average of the results in the 16 subbands with a threshold value.

[0050] A process of estimating the scale factor according
 10 to the subbands and extracting the watermark will be described below.

[0051] The estimation value g' of the scale factor is evaluated by an estimation method using the EM algorithm. The EM algorithm is used to estimate an average value μ_m of
 15 each component probability density function of a gaussian mixture model. The estimated rate g' is calculated through a linear regression analysis of the estimation value of μ_m , which is obtained by the EM algorithm. A variance σ_z^2 is updated using the rate g' . It is assumed that N number of
 20 observed values for estimation with respect to a positive integer N is $r_1, r_2, r_3, \dots, r_N$. A proposed estimation method consists of the repetition of the following steps. First, η_m and μ_m are calculated using equations 3 and 4.

$$[0052] \quad \eta_m^{(i)} = \frac{1}{N} \sum_{n=1}^N p(m | r_n, \theta^{(i-1)}), \text{ for } m=1,2,\dots,M \quad (\text{Eq. 3})$$

$$[0053] \quad \mu_m^{(i)} = \frac{\sum_{n=1}^N r_n p(m | r_n, \theta^{(i-1)})}{\sum_{n=1}^N p(m | r_n, \theta^{(i-1)})}, \text{ for } m=1,2,\dots,M \quad (\text{Eq. 4})$$

[0054] Here, the vector $\theta^{(i-1)}$ includes a value $\eta_m^{(i-1)}$, a value $\mu_m^{(i-1)}$, and a value $\sigma_z^{(i-1)}$ for $m=1,2,\dots,M$. Here, $p(m | r_n, \theta^{(i-1)})$ represents a posterior probability with respect to the coefficient $\theta^{(i-1)}$. Using the linear regression analysis, an estimation value $g^{(i)}$ of a rate with respect to the i -th repetition is calculated using a minimum value of a mean square error given by an equation 5.

$$[0055] \quad \sum_{m=1}^M \eta_m^{(i)} [\mu_m^{(i)} - g^{(i)} \mu_m^{(i-1)}]^2 \quad (\text{Eq. 5})$$

[0056] The estimation value $g^{(i)}$ of the rate is given by an equation 6.

$$[0057] \quad g^{(i)} = \frac{\sum_{m=1}^M \eta_m^{(i)} \mu_m^{(i)} \mu_m^{(i-1)}}{\sum_{m=1}^M \eta_m^{(i)} [\mu_m^{(i-1)}]^2} \quad (\text{Eq. 6})$$

[0058] The variance $\sigma_z^{(i-1)}$ is updated by an equation 7.

$$[0059] \quad \sigma_z^{(i)} = \sqrt{\frac{[g^{(i)}]^2 (D_2 - D_1)^2}{D_1} + (D_2 - D_1)} \quad (\text{Eq. 7})$$

[0060] In the proposed method, initial values of the coefficients are set like an equation 8.

$$[0061] \quad \sigma_z^{(0)} = \sqrt{\frac{(D_2 - D_1)^2}{D_1} + (D_2 - D_1)}$$

$$\mu_m^{(0)} = \frac{\Delta_e}{2} \left(m - \left[\frac{M-1}{2} \right] \right), \text{ for } m=1,2,\dots,M$$

$$\eta_m^{(0)} = \frac{1}{M}, \text{ for } m=1,2,\dots,M \quad (\text{Eq. 8})$$

[0062] The steps of updating these coefficients are repeated L times. A final rate is given by an equation 9.

$$[0063] \quad \hat{g} = g^L \quad (\text{Eq. 9})$$

[0064] The decoding process from the estimated rate g' is
5 achieved using the adjusted quantizer step size $\Delta_d = g' \Delta_e$.

[0065] A detecting process from the input signal r_n in each subband will be described below. First, the input signal r_n is quantized through the quantizer having the quantizer step size Δ_d and the dither $d(=0)$ to thereby provide a result
10 $Q_{\Delta_d,0}(r_n)$. Using g' that is a estimation result of g , the quantizer step size Δ_d of the decoder is made to have a value $g' \Delta_e$.

[0066] Assuming that \tilde{r}_n represents a quantization error, \tilde{r}_n is given by an equation 10.

$$15 \quad [0067] \quad \tilde{r}_n \cong r_n - Q_{\Delta_d,0}(r_n) \quad (\text{Eq. 10})$$

[0068] The estimated watermark signal \hat{d}_n is calculated by an equation 11.

$$[0069] \quad \hat{d}_n = \frac{4|\tilde{r}_n|}{\Delta_d} - 1 \quad (\text{Eq. 11})$$

[0070] An average of the results obtained in the 16
20 subbands is calculated, and a correlation between a resulting code and codes of a codebook is calculated. As a result, an index of code having the largest correlation is an embedded watermark information.

[0071] Referring to FIG. 6, the decoding method in the decoding apparatus includes the steps of: receiving an audio signal (601); dividing the audio signal into subbands (602); receiving a side information (603); estimating an scale factor from the side information and the audio signal by using an expectation maximization (EM) algorithm, and evaluating the quantizer step size from the estimated amplitude-scale rate (604); decoding a watermark from the subbands considering the evaluated quantizer step size (605); and summing up the decoded values to calculate an average, and calculating a correlation between the average and codes of a codebook to thereby obtain a watermark (606).

[0072] FIG. 7 illustrates simulation results when MP3 lossy compression and the amplitude scaling are applied, in which (A) is a case of no compression, (B) is a case of 192 kbps, and (C) is a case of 128 kbps.

[0073] In the graphs, the abscissa denotes a scale factor g and the ordinate denotes the bit error rate. A triangular solid line and a circular solid line represent a characteristic according to the prior art and the present invention, respectively. As shown in the graphs, although the bit error rate according to the prior art increases rapidly when the scale factor g increases, the bit error rate according to the present invention is not influenced by the amplitude scaling regardless of the scale factor.

[0074] As described above, according to the present invention, the scale factor is estimated from the watermarked signal itself without using additional signals such as a pilot signal. Therefore, even when an amplitude of the watermarked signal inputted into the decoder is changed, the watermark can be extracted without reducing an information embedding capacity. Additionally, the watermark signal is repeatedly embedded into areas of a low frequency subband, which is robust to a lossy compression or a low pass filtering, to areas of middle frequency subbands, which is robust to the amplitude scaling. Then, each result is summed up to extract the final watermark. Therefore, the present invention provides robustness in both the lossy compression and the amplitude scaling.

[0075] The lossy compression such as MP3 or the amplitude scaling of the audio signal may be used frequently in actual digital audio signal and considered as unintended attacks. In such case, the method and apparatus of the present invention is robust or resilient with respect to unintended changes, even when the watermarking is used for the purpose of embedding side information as well as protection of copyrights or a verification of integrity.

[0076] It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention. Thus, it is intended that the present invention covers the modifications and variations of this

invention provided they come within the scope of the appended claims and their equivalents.